

Multivariable Calculus

Thomas Kong

July 6, 2021

Contents

1	Preliminaries	5
1.1	Vector Spaces	5
1.2	Euclidean Space	10
1.3	Orthonormal Basis	13
1.4	Dual Space	17
1.5	Topology of Euclidean Space	21
1.6	Completeness	24
1.7	Continuity	24
1.8	Compactness	24
1.9	Equivalence of Norms	24
1.10	The Space $L(E, F)$	24

Chapter 1

Preliminaries

1.1 Vector Spaces

Our end goal is to do *multivariable* calculus. That is, we want to differentiate and integrate (real) vector-valued functions of several variables. To this end, we'll need to rigourously understand what a vector is so that we can comfortably manipulate them (beyond the highschool 'definition' of 'a vector is a quantity with both a magnitude and a direction'). Thus, we'll begin with a lightning review of basic concepts from linear algebra, beginning with the *actual* definition of a vector. A vector is simply an element of a structure called a *vector space*, which we define below:

Definition 1.1.1. A **vector space** over the field \mathbb{K} is a set V , together with the operations of 'vector addition' $f : V \times V \rightarrow V$ and 'scalar multiplication' $g : \mathbb{K} \times V \rightarrow V$, typically denoted by $f(x, y) = x + y$ and $g(\alpha, x) = \alpha x$, which satisfy the following axioms:

- (a) V is an abelian group with respect to the binary operation of vector addition.
- (b) Associativity of scalar multiplication: For every $v \in V$ and $a, b \in \mathbb{K}$, $(ab)v = a(bv)$.
- (c) Distributivity of scalar multiplication: For every $v \in V$ and $a, b \in \mathbb{K}$, $(a + b)v = av + bv$.
- (d) Distributivity of scalar multiplication over vector addition: For every $v, w \in V$ and $a \in \mathbb{K}$, $a(v + w) = av + aw$.
- (e) Multiplicative identity: For every $v \in V$, $1v = v$.

We will only consider vector spaces over the field of real numbers; from now on we will let the field \mathbb{K} be \mathbb{R} , and herein when we say 'vector space' we are really referring to a *real* vector space. The prototypical example of a vector space is the very first vector space that we all worked with before we even heard the term 'vector space': namely, n -tuples of real numbers, \mathbb{R}^n . Indeed, this will basically be the only vector space we will care about, as we will later see.

We'll now fly through some standard linear algebra concepts. Let V be a vector space. A **subspace** is a subset $S \subset V$ if S is itself a vector space under the same vector addition and scalar multiplication operations, restricted to S . Fortunately, to check that a subset of a vector space is a subspace, one does not actually have to check every single one of the vector space axioms. A necessary and sufficient condition for this is simply for the subset to be closed under vector addition and scalar multiplication, i.e for all $x, y \in S$ and $a \in \mathbb{R}$, $x + y \in S$ and $ax \in S$.

A mapping $f : S \rightarrow T$ between vector spaces that preserves the vector space structure is called a **linear map** or a **vector space homomorphism**, i.e for all $x, y \in S$ we have $f(x + y) = f(x) + f(y)$ and for all $a \in \mathbb{R}$ we have $f(ax) = af(x)$. A bijective homomorphism is called an **isomorphism**. Two vector spaces are *isomorphic* if there exists an isomorphism between them. Isomorphic vector spaces are essentially 'the same' for all intents and purposes; vector spaces that are isomorphic share all the same properties.

Let $f : S \rightarrow T$ be a vector space homomorphism. The set $\ker f = \{x \in S \mid f(x) = 0\}$ is called the **kernel** of f , and the set $\mathcal{R}(f) = \{f(x) \mid x \in S\}$ is called the **range** of f . It can be easily shown that $\ker f$ is a subspace of S and $\mathcal{R}(f)$ is a subspace of T . Another useful fact worth noting is that a linear map f is injective if and only if its kernel is trivial (i.e. $\ker f = \{0\}$). Hence, f is an isomorphism if and only if $\ker f = \{0\}$ and $\mathcal{R}(f) = T$.

A particularly important way that we use to classify vector spaces is the notion of dimension, which intuitively speaking is the ‘number of degrees of freedom’ it possesses. We will proceed to formalise this below. To find out how many degrees of freedom a vector space has, we essentially need to find the minimum number of fixed vectors required to write any arbitrary vector from the space as some weighted sum of these fixed vectors. We call such weighted sums of vectors **linear combinations**; a linear combination of the vectors in the subset $\{v_1, \dots, v_n\}$ is a sum $\sum_{i=1}^n c_i v_i$ for some constants $c_i \in \mathbb{R}$. The **span** of a subset $\{v_1, \dots, v_n\}$ is the set of all possible linear combinations of vectors in the set. Hence, in order to describe the entire vector space in terms of sums of vectors from one of its subsets $\{v_1, \dots, v_n\}$, we require that $V = \text{span}\{v_1, \dots, v_n\}$.

However, even if we are able to identify a finite subset $\{v_1, \dots, v_n\}$ of V such that $V = \text{span}\{v_1, \dots, v_n\}$, it is possible that we are able to find a smaller subset that does the trick, which suggests that the original subset may contain redundant information. The way we describe this redundancy is through the concept of linear dependence:

Definition 1.1.2. Let V be a vector space. A subset $A \subset V$ is **linearly independent** if for every finite subset $\{v_1, \dots, v_n\}$ of A , we have that $a_1 v_1 + \dots + a_n v_n = 0$ for scalars $a_1, \dots, a_n \in \mathbb{K}$ implies that $a_1 = \dots = a_n = 0$.

If A is not linearly independent, then it is **linearly dependent**.

At last, we arrive at our desired criteria for a subset of a vector space to summarise all of the information of V in the most minimalistic way possible:

Definition 1.1.3. Let V be a vector space. A **basis** for V is a linearly independent subset $A \subset V$ such that $V = \text{span } A$.

Fun remark: The definition we provide above is that of a *Hamel* basis. There are other types of bases. However, this distinction is not relevant for us in the finite dimensional setting, which we will be exclusively working in.

Given a basis $\{v_1, \dots, v_n\}$ for a vector space V , the representation of a vector $v \in V$ in that basis is unique. That is, if $v = a_1 v_1 + \dots + a_n v_n$ and $v = b_1 v_1 + \dots + b_n v_n$ for constants $a_i, b_i \in \mathbb{R}$ for $i = 1, \dots, n$, we have that $a_i = b_i$ for each i . This fact follows from the linear independence of the basis vectors.

So, we would like to say that the number of elements in a basis quantifies the number of degrees of freedom that the vector space possesses. However, one concern that arises is whether in our definition above that there is a possibility that there are bases with different numbers of elements. Fortunately, the answer is no, as we will now show:

Theorem 1.1.4. Let V be a vector space. Then every basis of V has the same number of elements, or are all infinite.

Proof. Suppose $\{v_1, \dots, v_n\}$ is a basis for V and suppose for a contradiction that $\{x_1, \dots, x_{n+1}\}$ is a linearly independent subset of V . We can write x_1 as a linear combination of the basis elements:

$$x_1 = a_{11}v_1 + \dots + a_{1n}v_n$$

where $a_{1i} \in \mathbb{R}$ for all $i = 1, \dots, n$. Since $\{x_1, \dots, x_{n+1}\}$ is a linearly independent set, then $x_1 \neq 0$, so that not all of the a_{1i} ’s are zero. Without loss of generality, suppose that $a_{11} \neq 0$ (else swap its label with one of the a_{1i} ’s that is nonzero, and also exchange the indices of the corresponding basis vectors accordingly). Then we can rearrange the above expression to solve for v_1 :

$$v_1 = \frac{1}{a_{11}}(x_1 - a_{12}v_2 - \dots - a_{1n}v_n)$$

It follows that the set $\{x_1, v_2, \dots, v_n\}$ is a basis for V . To see this, let $y \in V$. Then $y = c_1 v_1 + \dots + c_n v_n$ for constants $c_1, \dots, c_n \in \mathbb{R}$ (since the v_i 's form a basis for V). Thus, we have that:

$$y = \frac{c_1}{a_{11}} x_1 + \left(c_2 - \frac{a_{12}}{a_{11}} \right) v_2 + \dots + \left(c_n - \frac{a_{1n}}{a_{11}} \right) v_n$$

So y can be written as a linear combination of the vectors x_1, v_2, \dots, v_n , i.e. $y \in \text{span}\{x_1, v_2, \dots, v_n\}$. Since the choice of $y \in V$ was arbitrary, we have that $V = \text{span}\{x_1, v_2, \dots, v_n\}$. Linear independence of this set follows from the linear independence of the original basis. Indeed, consider the following vector equation:

$$k_1 x_1 + k_2 v_2 + \dots + k_n v_n = 0$$

for constants $k_1, \dots, k_n \in \mathbb{R}$. Expanding x_1 in terms of the original basis gives:

$$\begin{aligned} 0 &= k_1(a_{11}v_1 + \dots + a_{1n}v_n) + k_2v_2 + \dots + k_nv_n \\ &= k_1a_{11}v_1 + (k_1a_{12} + k_2)v_2 + \dots + (k_1a_{1n} + k_n)v_n \end{aligned}$$

Now, by linear independence of $\{v_1, \dots, v_n\}$, all of the coefficients in the above equation must vanish. In particular, we see that $k_1a_{11} = 0$, which implies that $k_1 = 0$ since $a_{11} \neq 0$ by assumption. Then, since $k_1a_{1i} + k_i = 0$ for all $i = 2, \dots, n$, we must have that $k_i = 0$ as well. Hence, $\{x_1, v_2, \dots, v_n\}$ is linearly independent, and our claim is proven.

What have we achieved? We have just replaced one of the basis vectors (namely v_1) with one of the vectors from the linearly independent set (namely x_1), and after the dust cleared we still have a basis for V . We will continue this process, gradually replacing all of the v_i 's with an x_i until we have a basis consisting of only x_i 's. We will prove that this process will work via induction.

Suppose we have replaced j of the basis vectors, and we have that $\{x_1, \dots, x_j, v_{j+1}, \dots, v_n\}$ is a basis for V (after possibly some relabelling of the vectors). We will show that we will obtain a basis by replacing one of the remaining v_i 's with x_{j+1} . The argument will follow quite similarly to our first replacement process. Write x_{j+1} in terms of the basis $\{x_1, \dots, x_j, v_{j+1}, \dots, v_n\}$:

$$x_{j+1} = a_{j+1,1}x_1 + \dots + a_{j+1,j}x_j + a_{j+1,j+1}v_{j+1} + \dots + a_{j+1,n}v_n$$

Since $x_{j+1} \neq 0$, then not all of the coefficients $a_{j+1,i}$ are zero. In fact, we must have that $a_{j+1,i} \neq 0$ for some $i \geq j+1$ (if this were not the case, then it follows that x_{j+1} is a linear combination of the vectors x_1, \dots, x_j , which contradicts the linear independence of the x_i 's). Without loss of generality, we'll take $a_{j+1,j+1} \neq 0$. Hence, we can write:

$$v_{j+1} = \frac{1}{a_{j+1,j+1}}(x_{j+1} - a_{j+1,1}x_1 - \dots - a_{j+1,j}x_j - a_{j+1,j+2}v_{j+2} - \dots - a_{j+1,n}v_n)$$

We'll now demonstrate that $\{x_1, \dots, x_{j+1}, v_{j+2}, \dots, v_n\}$ is a basis for V . Let $y \in V$. Writing y in terms of the basis $\{x_1, \dots, x_j, v_{j+1}, \dots, v_n\}$ yields:

$$y = c_1x_1 + \dots + c_jx_j + c_{j+1}v_{j+1} + \dots + c_nv_n$$

for some constants $c_1, \dots, c_n \in \mathbb{R}$. Substituting in our expression for v_{j+1} gives:

$$y = \left(c_1 - \frac{a_{j+1,1}}{a_{j+1,j+1}} \right) x_1 + \dots + \left(c_j - \frac{a_{j+1,j}}{a_{j+1,j+1}} \right) x_j + \frac{c_{j+1}}{a_{j+1,j+1}} x_{j+1} + \left(c_{j+2} - \frac{a_{j+1,j+2}}{a_{j+1,j+1}} \right) v_{j+2} + \dots + \left(c_n - \frac{a_{j+1,n}}{a_{j+1,j+1}} \right) v_n$$

Hence $y \in \text{span}\{x_1, \dots, x_{j+1}, v_{j+2}, \dots, v_n\}$. Since the choice of $y \in V$ was arbitrary, it follows that $V = \text{span}\{x_1, \dots, x_{j+1}, v_{j+2}, \dots, v_n\}$.

Now we'll demonstrate linear independence. Consider the following vector equation:

$$k_1x_1 + \dots + k_{j+1}x_{j+1} + k_{j+2}v_{j+2} + \dots + k_nv_n = 0$$

for constants $k_1, \dots, k_n \in \mathbb{R}$. Expanding x_{j+1} in terms of the basis $\{x_1, \dots, x_j, v_{j+1}, \dots, v_n\}$ yields:

$$\begin{aligned} 0 &= k_1 x_1 + \dots + k_{j+1}(a_{j+1,1}x_1 + \dots + a_{j+1,j}x_j + a_{j+1,j+1}v_{j+1} + \dots + a_{j+1,n}v_n) + k_{j+2}v_{j+2} + \dots + k_n v_n \\ &= (k_1 + k_{j+1}a_{j+1,1})x_1 + \dots + (k_j + k_{j+1}a_{j+1,j})x_j + k_{j+1}a_{j+1,j+1}v_{j+1} + (k_{j+2} + k_{j+1}a_{j+1,j+2})v_{j+2} + \\ &\quad \dots + (k_n + k_{j+1}a_{j+1,n})v_n \end{aligned}$$

By linear independence of $\{x_1, \dots, x_j, v_{j+1}, \dots, v_n\}$, all of the above coefficients must vanish. In particular, consider the $j+1$ coefficient: $k_{j+1}a_{j+1,j+1} = 0$. From this, we conclude that $k_{j+1} = 0$ since $a_{j+1,j+1} \neq 0$, and hence looking at the remaining coefficients we must have $k_i = 0$ for all $i = 1, \dots, n$, thus demonstrating the linear independence of $\{x_1, \dots, x_{j+1}, v_{j+2}, \dots, v_n\}$, and hence proving that it is a basis for V .

By induction on the finite set $\{1, \dots, n\}$, we can carry out our replacement procedure and end up with the fact that $\{x_1, \dots, x_n\}$ forms a basis for V . Now, since $x_{n+1} \in V$, we can write it as a nonzero linear combination of the basis vectors x_1, \dots, x_n . However, this contradicts the linear independence of the set $\{x_1, \dots, x_{n+1}\}$. We conclude that no linearly independent set can have any more vectors than any basis of V , if there exists a finite basis for V . Since bases are linearly independent sets, then no finite basis can have any more vectors than any other basis. So if there exists a finite basis for V , then every basis for V must also be finite and have the same number of elements. The only other possibility is that every basis of V contains infinitely many elements. \square

This justifies the following definition:

Definition 1.1.5. Let V be a vector space. The **dimension** of V , denoted $\dim V$, is the number of elements in any basis of V if they are finite, or ∞ otherwise.

We will only consider finite dimensional vector spaces, that is, vector spaces in which there exists a basis with finitely many elements. Doing calculus on infinite dimensional vector spaces crosses into the realm of functional analysis, but we won't stray in that direction in the scope of these notes.

One key advantage of working in finite dimensional spaces is that things all turn out to be significantly simpler. One easy consequence of the previous theorem is that in an n -dimensional vector space, if we have a linearly independent set consisting of n elements, it must actually span the entire space and is thus automatically a basis.

Corollary 1.1.6. Let V be an n -dimensional vector space and let $\{v_1, \dots, v_n\}$ be a linearly independent set. Then $\{v_1, \dots, v_n\}$ is a basis for V .

Proof. Suppose for a contradiction that $\text{span}\{v_1, \dots, v_n\} \neq V$. So there exists a $x \in V$ such that x is not a linear combination of the v_i 's. Hence, $\{v_1, \dots, v_n, x\}$ is a linearly independent set. So we have a linearly independent set consisting of $n+1$ elements in an n -dimensional space. This is impossible, since every basis of V must contain n elements and no linearly independent set can have more elements than any basis of V . We must conclude that $V = \text{span}\{v_1, \dots, v_n\}$, and hence $\{v_1, \dots, v_n\}$ is a basis for V . \square

Another simplification in the finite dimensional setting is that there actually aren't 'that many' different types of finite dimensional vector spaces for each dimension $n \in \mathbb{N}$ to study; we can easily completely classify every real finite dimensional vector space up to isomorphism - even better: they are all structurally identical to \mathbb{R}^n !

Theorem 1.1.7. Let V be an n -dimensional vector space. Then V is isomorphic to \mathbb{R}^n .

Proof. Let $\{v_1, \dots, v_n\}$ be a basis for V . Define the mapping $\phi: V \rightarrow \mathbb{R}^n$ to act on $v = a_1 v_1 + \dots + a_n v_n$ by:

$$\begin{aligned} \phi(v) &= \phi(a_1 v_1 + \dots + a_n v_n) \\ &= (a_1, \dots, a_n)^T \end{aligned}$$

So ϕ is the canonical map that maps a vector $v \in V$ to an n -tuple of real numbers, which are simply the coordinates of v with respect to a certain basis of V . We will show that ϕ is an isomorphism.

Note that the elements mapped by ϕ to $0 \in \mathbb{R}^n$ must have zero coefficients when expanded in terms of the basis $\{v_1, \dots, v_n\}$, and thus can only be the zero vector. This shows that $\ker \phi = \{0\}$ and hence that ϕ is injective.

Let $x \in \mathbb{R}^n$. Then $x = (a_1, \dots, a_n)^T$ for some $a_1, \dots, a_n \in \mathbb{R}$. Thus, it is easy to see that the element $v = a_1v_1 + \dots + a_nv_n \in V$ is mapped to x by ϕ . Since the choice of $x \in \mathbb{R}^n$ was arbitrary, it follows that ϕ is surjective.

Let $v = a_1v_1 + \dots + a_nv_n \in V$ and $w = b_1v_1 + \dots + b_nv_n \in V$. Then:

$$\begin{aligned}\phi(v + w) &= (a_1 + b_1, \dots, a_n + b_n) \\ &= (a_1, \dots, a_n) + (b_1, \dots, b_n) \\ &= \phi(v) + \phi(w)\end{aligned}$$

Let $v \in V$ be as above and let $k \in \mathbb{R}$. Then:

$$\begin{aligned}\phi(kv) &= (ka_1, \dots, ka_n) \\ &= k(a_1, \dots, a_n) \\ &= k\phi(v)\end{aligned}$$

Thus, ϕ is a vector space homomorphism. Together with the fact that ϕ is bijective, we have hence shown that ϕ is an isomorphism. \square

So any two n -dimensional vector spaces are isomorphic. This is what we meant when we said that the only n -dimensional vector space we will care about is \mathbb{R}^n - all other instances of an n -dimensional vector space are algebraically equivalent, so we may as well just consider this simple example as being *the* n -dimensional vector space.

One more tidbit that we should address is to show that two vector spaces of different dimensions cannot be isomorphic, so that finite dimensional vector spaces are isomorphic if and only if their dimensions are equal (i.e dimension is a vector space property that is preserved by isomorphism). This fact would imply that we have completely classified every single finite dimensional vector space: there is only one unique vector space for each dimension $n \in \mathbb{N}$ - namely \mathbb{R}^n , and $\mathbb{R}^n \simeq \mathbb{R}^m$ if and only if $m = n$.

Proposition 1.1.8. *Let V and W be finite dimensional vector spaces. If V and W are isomorphic, then $\dim V = \dim W$.*

Proof. Let $n = \dim V$ and let $\{v_1, \dots, v_n\}$ be a basis for V . Let $\phi : V \rightarrow W$ be a vector space isomorphism. We will show that $\{\phi(v_1), \dots, \phi(v_n)\}$ is a basis for W .

First, we prove linear independence. Consider the following vector equation:

$$a_1\phi(v_1) + \dots + a_n\phi(v_n) = 0$$

where $a_i \in \mathbb{R}$ for $i = 1, \dots, n$. By linearity of the mapping ϕ , this is equivalent to:

$$\phi(a_1v_1 + \dots + a_nv_n) = 0$$

So $a_1v_1 + \dots + a_nv_n \in \ker \phi$. However, since ϕ is injective, $\ker \phi = \{0\}$. It follows that:

$$a_1v_1 + \dots + a_nv_n = 0$$

By linear independence of $\{v_1, \dots, v_n\}$, it follows that $a_i = 0$ for all $i = 1, \dots, n$. Hence, $\{\phi(v_1), \dots, \phi(v_n)\}$ is linearly independent.

Let $w \in W$. Since ϕ is surjective, there exists a $v \in V$ such that $w = \phi(v)$. Express v in terms of the basis $\{v_1, \dots, v_n\}$:

$$v = c_1v_1 + \dots + c_nv_n$$

Then:

$$\begin{aligned} w &= \phi(v) \\ &= \phi(c_1v_1 + \cdots + c_nv_n) \\ &= c_1\phi(v_1) + \cdots + c_n\phi(v_n) \end{aligned} \quad (\text{by linearity of } \phi)$$

Hence, $w \in \text{span}\{\phi(v_1), \dots, \phi(v_n)\}$. Since the choice of $w \in W$ is arbitrary, it follows that $W = \text{span}\{\phi(v_1), \dots, \phi(v_n)\}$. This shows that $\{\phi(v_1), \dots, \phi(v_n)\}$ is a basis for W , which has n elements. By definition, this implies that $\dim W = n = \dim V$. \square

In summary, we have proven the following very important result about finite dimensional vector spaces:

Theorem 1.1.9. *Let V and W be finite dimensional vector spaces. Then V and W are isomorphic if and only if $\dim V = \dim W$.*

1.2 Euclidean Space

Now that we have formulated all of the algebraic structure of the multidimensional spaces we will be working in, we will need to endow some further structure onto these spaces in order to do calculus. Of great importance is the ability to measure distances between points/vectors in our space. At the heart of analysis/calculus is the notion of *limits* and *convergence*. We will also frequently want to find bounds and estimates for quantities that may very well be vector-valued. So what we ultimately want is to prescribe some notion of ‘length’ for vectors, which will in turn naturally define a distance between two vectors (simply take the length of the difference of the two vectors). This is captured by the concept of a vector *norm*.

Definition 1.2.1. Let V be a vector space. A norm on V is a function $\|-\| : V \rightarrow \mathbb{R}$ satisfying:

- (a) Positivity: For all $x \in V$ $\|x\| \geq 0$. Furthermore, $\|x\| = 0$ if and only if $x = 0$.
- (b) Homogeneity of degree 1: For all $x \in V$ and $a \in \mathbb{R}$, $\|ax\| = |a|\|x\|$.
- (c) Triangle inequality: for all $x, y \in V$, $\|x + y\| \leq \|x\| + \|y\|$.

A vector space together with a norm is called a **normed vector space**.

Another useful concept from coordinate geometry is the notion of angles. This will lend us the ability to define properties such as orthogonality, and perform operations such as projecting a vector onto a subspace. This structure is given rise to by an *inner product*, which we define below:

Definition 1.2.2. Let V be a vector space. A (real) **inner product** on V is a function $\langle -, - \rangle : V \times V \rightarrow \mathbb{R}$ satisfying:

- (a) Symmetry: for all $x, y \in V$, $\langle x, y \rangle = \langle y, x \rangle$.
- (b) Distributivity: for all $x, y, z \in V$, $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.
- (c) Homogeneity of degree 1: For all $x, y \in V$ and $a \in \mathbb{R}$, $\langle ax, y \rangle = a \langle x, y \rangle$.
- (d) Positivity: For all $x \in V$, $\langle x, x \rangle \geq 0$. Moreover, $\langle x, x \rangle = 0$ if and only if $x = 0$.

A vector space together with an inner product is called an **inner product space**.

By symmetry of the inner product, we can deduce from the definition that $\langle x, ay \rangle = a \langle x, y \rangle$ and $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$. Hence, an inner product on a vector space is simply a bilinear function (i.e it becomes a linear function of one variable if we fix the value of the other input variable).

You are likely already quite familiar with an inner product on \mathbb{R}^n from coordinate geometry, more commonly known as the dot product or scalar product. Given elements $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ and $y = (y_1, \dots, y_n)^T \in \mathbb{R}^n$, we define the dot product as:

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i$$

The notation $x \cdot y$ is more commonly used instead of $\langle x, y \rangle$ in this context. It is not too hard to see that the dot product is indeed an inner product on \mathbb{R}^n .

One powerful fact about inner products is that they can be used to define a norm in a very natural way. Let V be an inner product space. Then for any $x \in V$, we set the norm of x to be given by $\|x\| = \langle x, x \rangle^{\frac{1}{2}}$. We say that the inner product ‘induces’ a norm on V . So in some sense, once we had defined the notion of angles on our vector space, we have already implicitly fixed some compatible notion of distance as well for free. The positivity and homogeneity properties of the norm follow readily from the properties of the inner product and from the definition of the induced norm. Proving the triangle inequality will take a little more work, and will rely on an identity known as the *Schwarz inequality*:

Theorem 1.2.3. (*Schwarz Inequality*) *Let V be an inner product space. For every $x, y \in V$, we have:*

$$\langle x, y \rangle \leq \|x\| \|y\|$$

Proof. Let $x, y \in V$ and let $t \in \mathbb{R}$. By positivity of the inner product, we have that $\langle x - ty, x - ty \rangle \geq 0$. This can be rewritten as:

$$\begin{aligned} 0 &\leq \langle x - ty, x - ty \rangle \\ &= \langle x, x \rangle - t \langle x, y \rangle - t \langle y, x \rangle + t^2 \langle y, y \rangle \\ &= \|x\|^2 - 2t \langle x, y \rangle + t^2 \|y\|^2 \end{aligned}$$

So we have a real quadratic equation in the variable t . The quadratic can have at most a single zero, since $\langle x - ty, x - ty \rangle = 0$ if and only if $x - ty = 0$, i.e. $x = ty$. Hence, the discriminant of the quadratic must be either 0 (which corresponds to the single zero) or negative (no real zeros). The discriminant of the quadratic is:

$$4 \langle x, y \rangle^2 - 4 \|x\|^2 \|y\|^2$$

Putting this all together, we conclude that:

$$\begin{aligned} 4 \langle x, y \rangle^2 - 4 \|x\|^2 \|y\|^2 &\leq 0 \\ \langle x, y \rangle &\leq \|x\| \|y\| \end{aligned}$$

□

From the above proof, we note that equality occurs in the Schwarz inequality if and only if $x = ty$ (i.e. the vectors are scalar multiples of each other).

We are now ready to tackle the proof that the inner product gives rise to a vector norm:

Theorem 1.2.4. *Let V be an inner product space. The function $\|-\| : V \rightarrow \mathbb{R}$ defined by $\|x\| = \langle x, x \rangle^{\frac{1}{2}}$ is a norm on V .*

Proof. Positivity: Let $x \in V$. We have that $\|x\|^2 = \langle x, x \rangle \geq 0$, and hence $\|x\| \geq 0$. Now, from $\|x\|^2 = \langle x, x \rangle$, we can easily see that $\|x\| = 0$ if and only if $x = 0$, by positivity of the inner product.

Homogeneity of degree 1: Let $x \in V$ and $a \in \mathbb{R}$. Then $\|ax\|^2 = \langle ax, ax \rangle = a^2 \langle x, x \rangle = a^2 \|x\|^2$. Taking square roots gives $\|ax\| = |a| \|x\|$.

Triangle inequality: Let $x, y \in V$. Then:

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + 2 \langle x, y \rangle + \langle y, y \rangle \\ &= \|x\|^2 + 2 \langle x, y \rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 && \text{(by the Schwarz inequality)} \\ &= (\|x\| + \|y\|)^2 \end{aligned}$$

Taking square roots gives the desired result: $\|x + y\| \leq \|x\| + \|y\|$.

□

Finally, we need to formalise the notion of measuring distance in our vector space. The idea of distance is captured by a ‘metric’, which intuitively speaking is a distance-measuring function.

Definition 1.2.5. Let X be a set. A **metric** on X is a function $d : X \times X \rightarrow \mathbb{R}$ satisfying:

- (a) Positivity: for all $x, y \in X$, $d(x, y) \geq 0$. Moreover, $d(x, y) = 0$ if and only if $x = y$.
- (b) Symmetry: for all $x, y \in X$, $d(x, y) = d(y, x)$.
- (c) Triangle inequality: for all $x, y, z \in X$, $d(x, z) \leq d(x, y) + d(y, z)$.

A set together with a metric is called a **metric space**.

We mentioned earlier that once we had a norm on a vector space, this was sufficient to define a notion of distance. Specifically, we claimed that to measure the distance between two vectors x and y , we simply need to compute the length of the vector $x - y$ using the vector norm. Thus, we claim that the norm naturally induces a metric on the vector space.

Theorem 1.2.6. Let V be a normed vector space. The function $d : V \times V \rightarrow \mathbb{R}$ defined by $d(x, y) = \|x - y\|$ is a metric on V .

Proof. Positivity: Let $x, y \in V$. Then $d(x, y) = \|x - y\| \geq 0$ by positivity of the norm. Furthermore, $d(x, y) = 0$ if and only if $x - y = 0$, i.e $x = y$.

Symmetry Let $x, y \in V$. Then $d(x, y) = \|x - y\| = \|(y - x)\| = |-1|\|y - x\| = d(y, x)$, which follows from the homogeneity of the norm.

Triangle inequality: Let $x, y, z \in V$. Then:

$$\begin{aligned} d(x, z) &= \|x - z\| \\ &= \|(x - y) + (y - z)\| \\ &\leq \|x - y\| + \|y - z\| && \text{(since the norm obeys the triangle inequality)} \\ &= d(x, y) + d(y, z) \end{aligned}$$

□

In summary, by defining an inner product on our vector space, the inner product induces a norm, which in turn induces a metric. So we get the geometric concepts of angle, length and distance all at once via an inner product. Hence, we call an n -dimensional vector space with an inner product **Euclidean n -space**; where ‘Euclidean’ refers to the fact that we have a means to measure distances in our space.

Returning to the example of \mathbb{R}^n , we see that \mathbb{R}^n together with the dot product defines a Euclidean n -space. In fact, it is interesting to note that the metric induced by the dot product is simply the function that returns the Euclidean distance between two points in \mathbb{R}^n (i.e the Euclidean metric d_2):

$$d_2(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

In the previous section, we showed that all n -dimensional vector spaces are isomorphic to each other, and in particular, \mathbb{R}^n , so we can just talk about \mathbb{R}^n without any loss of generality. Even better, it is the case that all n -dimensional inner product spaces are isomorphic, and moreover the inner product is preserved by the isomorphism. Thus, there is essentially only one Euclidean n -space up to isomorphism: namely, \mathbb{R}^n with the dot product. The proof of this will be the subject of the following section.

1.3 Orthonormal Basis

Most of the time when we are working with vector spaces, it is very useful to pick a basis and represent all vectors in terms of that basis. Some bases are more convenient to work with than others. Indeed, in a Euclidean space, now that we have some notion of angles provided by the inner product, we can define what it means for two vectors to be orthogonal (the generalisation of perpendicularity from coordinate geometry). Hence, we could try to find bases that are *orthonormal*; the basis vectors are orthogonal and normalised. We'll define these concepts below:

Definition 1.3.1. Let E be a Euclidean n -space. Two elements $x, y \in E$ are **orthogonal** if $\langle x, y \rangle = 0$.

A subset $A \subset E$ is said to be orthogonal if every pair of distinct elements from A are orthogonal.

Definition 1.3.2. Let E be a Euclidean n -space. A set $A \subset E$ is **orthonormal** if A is orthogonal and $\|x\| = 1$ for all $x \in A$.

Orthonormality is a powerful condition. One immediate consequence of the orthonormality of a set is linear independence:

Proposition 1.3.3. *Every orthonormal set in a Euclidean n -space is linearly independent.*

Proof. Suppose $\{e_1, \dots, e_m\}$ is an orthonormal set in a Euclidean n -space E ($m \leq n$). Consider the vector equation $a_1 e_1 + \dots + a_m e_m = 0$ for constants $a_1, \dots, a_m \in \mathbb{R}$. Taking the inner product of e_i , $i = 1, \dots, m$ with the above equation yields:

$$\begin{aligned} 0 &= \left\langle e_i, \sum_{j=1}^m a_j e_j \right\rangle \\ &= \sum_{j=1}^m a_j \langle e_i, e_j \rangle \\ &= \sum_{j=1}^m a_j \delta_{ij} \\ &= a_i \end{aligned}$$

where we have made use of the Kronecker delta symbol δ_{ij} , which returns 1 if $i = j$ and 0 otherwise.

So $a_i = 0$ for all $i = 1, \dots, m$. This proves the linear independence of the finite set $\{e_1, \dots, e_m\}$.

Now suppose for a contradiction there exists an orthonormal set $A \subset E$ with more than n elements. Pick n elements from A to form the orthonormal set $\{e_1, \dots, e_n\}$, which is linearly independent from what we have just proven. Hence, it forms a basis for E . Pick another *distinct* element $e_{n+1} \in A$, and write it as a linear combination of the basis elements:

$$e_{n+1} = k_1 e_1 + \dots + k_n e_n$$

Now take the inner product of the above equation with e_{n+1} :

$$\begin{aligned} \langle e_{n+1}, e_{n+1} \rangle &= \left\langle e_{n+1}, \sum_{j=1}^n k_j e_j \right\rangle \\ &= \sum_{j=1}^n k_j \langle e_{n+1}, e_j \rangle \\ &= 0 \end{aligned}$$

However, $\langle e_{n+1}, e_{n+1} \rangle = \|e_{n+1}\|^2 = 1$, so we have that $1 = 0$; a contradiction. It follows that there cannot exist an orthonormal set in E with more than n elements. Since we have proven that every orthonormal set with $m \leq n$ elements is linearly independent, then we thus conclude that every orthonormal set in E is linearly independent. \square

One interesting consequence of the above proof is the following. Suppose we could write a vector x from a Euclidean space E as a linear combination of orthonormal vectors $x = a_1e_1 + \cdots + a_ne_n$. Then we can define a mapping $\pi_i : E \rightarrow \mathbb{R}$ for $i = 1, \dots, n$ by $\pi_i(x) = \langle x, e_i \rangle = a_i$. This illustrates another advantage of using an orthonormal basis to represent vectors: finding the coefficients becomes a straightforward task - one simply needs to calculate $\pi_i(x)$ to find the i^{th} coefficient. We call the mapping π_i a *projection operator*; essentially π_i is projecting the vector x onto the subspace spanned by the i^{th} basis vector and picking out the amplitude of the vector projection.

Now that we are sufficiently hyped up about orthonormal bases, the next natural question to ask is whether it is always possible to find an orthonormal basis for any Euclidean n -space. We know from the previous result that every orthonormal set is linearly independent. Combined with the fact that in an n -dimensional space, any linearly independent set consisting of n elements is automatically a basis, all we need to do is prove the existence of an orthonormal set consisting of n elements. Spoiler alert: yes, this is certainly possible. Better yet, the proof of this result will be constructive, it illustrates a standard algorithm called the ‘Gram-Schmidt’ process which transforms an arbitrary basis into an orthonormal basis.

Proposition 1.3.4. *Let E be a Euclidean n -space. Then there exists an orthonormal subset $A \subset E$ consisting of n elements.*

Proof. Let $\{v_1, \dots, v_n\}$ be a linearly independent set in E . It is worth pointing out that none of the vectors are 0 due to linear independence. Thus, we can create an orthonormal set of one element by taking the first element from the set and normalising it: define $e_1 = v_1/\|v_1\|$. Then $\{e_1\}$ is an orthonormal set, and e.g. $\{e_1, v_2\}$ is a linearly independent set (since e_1 is simply a scalar multiple of v_1). This sets us up to prove the inductive step.

Suppose that we had that $\{e_1, \dots, e_k\}$ is an orthonormal set ($k < n$) and that $\{e_1, \dots, e_k, v_{k+1}\}$ is a linearly independent set. Define:

$$y_{k+1} = v_{k+1} - \sum_{j=1}^k \langle v_{k+1}, e_j \rangle e_j$$

It follows that $y_{k+1} \neq 0$ since y_{k+1} is a linear combination of e_1, \dots, e_k, v_{k+1} , which is a linearly independent set - and so the only way for such a linear combination to yield the zero vector is if all of the coefficients are 0, which is not the case here. Hence, $\|y_{k+1}\| \neq 0$, and so we can safely define:

$$e_{k+1} = \frac{y_{k+1}}{\|y_{k+1}\|}$$

So $\|e_{k+1}\| = 1$, and for all $i = 1, \dots, k$:

$$\begin{aligned} \langle e_{k+1}, e_i \rangle &= \left\langle \frac{1}{\|y_{k+1}\|} \left(v_{k+1} - \sum_{j=1}^k \langle v_{k+1}, e_j \rangle e_j \right), e_i \right\rangle \\ &= \frac{1}{\|y_{k+1}\|} \left(\langle v_{k+1}, e_i \rangle - \sum_{j=1}^k \langle v_{k+1}, e_i \rangle \langle e_j, e_i \rangle \right) \\ &= \frac{1}{\|y_{k+1}\|} \left(\langle v_{k+1}, e_i \rangle - \sum_{j=1}^k \langle v_{k+1}, e_j \rangle \delta_{ij} \right) \\ &= \frac{1}{\|y_{k+1}\|} (\langle v_{k+1}, e_i \rangle - \langle v_{k+1}, e_i \rangle) \\ &= 0 \end{aligned}$$

Hence, $\{e_1, \dots, e_{k+1}\}$ is an orthonormal set. If $k+1 < n$, then there is still an element v_{k+2} that cannot be written as a linear combination of the orthonormal set $\{e_1, \dots, e_{k+1}\}$, since the e_i ’s are simply linear combinations of v_1, \dots, v_{k+1} , but $\{v_1, \dots, v_{k+2}\}$ is a linearly independent set. It follows that $\{e_1, \dots, e_{k+1}, v_{k+2}\}$ is a linearly independent set. We can then proceed via induction to deduce that we can produce an orthonormal set $\{e_1, \dots, e_n\}$. \square

The previous two results together imply the following:

Theorem 1.3.5. *Every Euclidean n -space has an orthonormal basis.*

Let us return to the original problem that we wished to address in this section: that there is only one Euclidean n -space up to isomorphism. Since we have already proven that the vector space structure of two Euclidean n -spaces are equivalent, we simply need to prove that the ‘extra structure’ of the Euclidean space is also preserved under a vector space isomorphism. Let’s first make explicit what ‘extra stuff’ needs to be preserved:

Definition 1.3.6. Let E and F be Euclidean spaces. We say that E and F are **isomorphic** if there exists a mapping $\phi : E \rightarrow F$ such that ϕ is a vector space isomorphism that also preserves the inner product, i.e. for all $x, y \in E$, $\langle x, y \rangle = \langle \phi(x), \phi(y) \rangle$.

So we require that just the inner product needs to be preserved by the isomorphism. But what about the corresponding norm and metric? Well, if the inner product between Euclidean spaces E and F is preserved by an isomorphism $\phi : E \rightarrow F$, it follows that ϕ also preserves the induced norms and metrics on these spaces. To see this, let $x \in E$ and notice that $\|x\|^2 = \langle x, x \rangle = \langle \phi(x), \phi(x) \rangle = \|\phi(x)\|^2$. Now that we have shown that ϕ is norm-preserving, the fact that distances are preserved follows easily. For any $x, y \in E$ we have that $d(x, y) = \|x - y\| = \|\phi(x - y)\| = \|\phi(x) - \phi(y)\| = d(\phi(x), \phi(y))$. Hence, we see that ϕ is an isometry (a distance preserving function).

Recall from our proof that every n -dimensional vector space is isomorphic (to \mathbb{R}^n) (Theorem 1.1.7), we constructed a function that essentially mapped one basis onto another. In order to preserve the inner product of the Euclidean spaces, the isomorphisms that we will construct in the following proof will be a subset of the isomorphisms that did the trick in the more general vector space case. Specifically, the isomorphisms between Euclidean spaces will, loosely speaking, send *orthonormal bases to orthonormal bases*.

Theorem 1.3.7. *Let E and F be Euclidean n -spaces. Then E and F are isomorphic.*

Proof. Let $\{e_1, \dots, e_n\}$ be an orthonormal basis in E , and $\{\bar{e}_1, \dots, \bar{e}_n\}$ an orthonormal basis in F (we know that orthonormal bases for Euclidean spaces exist by the previous theorem which we have just proven). We can express any $x \in E$ as a linear combination of the basis vectors: $x = a_1e_1 + \dots + a_ne_n$. We can hence define the map $\phi : E \rightarrow F$ by:

$$\phi(x) = a_1\bar{e}_1 + \dots + a_n\bar{e}_n$$

The form of ϕ is exactly that of the isomorphism which we have constructed in the proof of Theorem 1.1.7, i.e. a map that sends a vector represented in a given basis to a vector in the other space with the same coordinates/coefficients but with respect to a basis in the target space¹. Thus, we already have that ϕ is a vector space isomorphism. We thus simply need to demonstrate that ϕ preserves the inner product. Let $x \in E$ as above and let $y = b_1e_1 + \dots + b_ne_n \in E$. Then:

$$\begin{aligned} \langle x, y \rangle &= \left\langle \sum_{i=1}^n a_i e_i, \sum_{j=1}^n b_j e_j \right\rangle \\ &= \sum_{i=1}^n a_i \left\langle e_i, \sum_{j=1}^n b_j e_j \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i b_j \langle e_i, e_j \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i b_j \delta_{ij} && \text{(by orthonormality of the basis vectors)} \\ &= \sum_{i=1}^n a_i b_i \end{aligned}$$

¹Specifically, we mapped a basis from an n -dimensional vector space to the standard basis in \mathbb{R}^n in that proof.

But we also have that:

$$\begin{aligned}
 \langle \phi(x), \phi(y) \rangle &= \left\langle \sum_{i=1}^n a_i \bar{e}_i, \sum_{j=1}^n b_j \bar{e}_j \right\rangle \\
 &= \sum_{i=1}^n \sum_{j=1}^n a_i b_j \langle \bar{e}_i, \bar{e}_j \rangle \\
 &= \sum_{i=1}^n \sum_{j=1}^n a_i b_j \delta_{ij} && \text{(by orthonormality of the basis vectors)} \\
 &= \sum_{i=1}^n a_i b_i
 \end{aligned}$$

So $\langle x, y \rangle = \langle \phi(x), \phi(y) \rangle$, and so the inner product is preserved. Hence, ϕ is an isomorphism of Euclidean spaces. \square

Hence, every Euclidean n -space is ‘the same’ from a structural point of view, so we might as well work with the easiest example of it we can think of: i.e \mathbb{R}^n with the dot product.

While on the topic of orthonormal bases, we’ll conclude this section with a fun digression. Let E be a Euclidean n -space, and suppose we had two orthonormal bases in E : $\{e_1, \dots, e_n\}$ and $\{\bar{e}_1, \dots, \bar{e}_n\}$. We can write the barred basis vectors \bar{e}_i in terms of the basis vectors e_j :

$$\bar{e}_i = \sum_{j=1}^n a_{ij} e_j \quad i = 1, \dots, n$$

The n^2 constants $a_{ij} \in \mathbb{R}$ form a $n \times n$ matrix O , which represents the linear map $\phi : E \rightarrow E$ defined by $\phi(a_1 e_1 + \dots + a_n e_n) = a_1 \bar{e}_1 + \dots + a_n \bar{e}_n$, which we know from above to be an automorphism of Euclidean spaces. So if we wrote the components of a vector $x \in E$ with respect to the orthonormal basis $\{e_1, \dots, e_n\}$ as a column vector, then $\phi(x) = O x$, where the components of $\phi(x)$ are also expressed in a column vector. What properties does the matrix O have? Well, since $\langle \bar{e}_i, \bar{e}_j \rangle = \delta_{ij}$ by orthonormality, we can substitute our expression for the barred basis vectors in terms of the unbarred basis vectors to obtain:

$$\begin{aligned}
 \delta_{ij} &= \left\langle \sum_{k=1}^n a_{ik} e_k, \sum_{l=1}^n a_{jl} e_l \right\rangle \\
 &= \sum_{k=1}^n \sum_{l=1}^n a_{ik} a_{jl} \langle e_k, e_l \rangle \\
 &= \sum_{k=1}^n \sum_{l=1}^n a_{ik} a_{jl} \delta_{kl} \\
 &= \sum_{k=1}^n a_{ik} a_{jk}
 \end{aligned}$$

Suppose that b_{ij} are the components of the matrix O^T (the matrix transpose of O). Then $b_{kj} = a_{jk}$, so that we have that $\delta_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$. Since δ_{ij} are the components of the $n \times n$ identity matrix I_n , we can write the above equation in the following matrix form:

$$I_n = O O^T$$

Since O is the matrix representation of a vector space automorphism, it is invertible, and hence by premultiplying both sides of the above equation by O^{-1} , we arrive at:

$$O^{-1} = O^T$$

i.e the matrix O is an orthogonal matrix. This reveals that orthogonal transformations on Euclidean spaces are simply the ones that map orthonormal bases to orthonormal bases.

1.4 Dual Space

In this section, we study a very special type of vector space mapping. Namely, the linear maps that take a vector as an input and returns a scalar. Such maps are called **linear functionals**.

Definition 1.4.1. Let V be an n -dimensional vector space. A **linear functional** on V is a linear mapping $f : V \rightarrow \mathbb{R}$.

Definition 1.4.2. Let V be an n -dimensional vector space. The **dual space** V^* is the set of all linear functionals on V .

The dual space of a vector space is itself a vector space under the operations of addition and scalar multiplication of real-valued functions. Elements of the dual space are sometimes also referred to as *covectors*.

There are many important linear functionals that are central to analysis. To name a few: the vector norm $\|\cdot\|$, integral operators, scalar potentials, density functions. It is therefore quite valuable to study them in a general setting.

In the finite dimensional setting, the study of linear functionals and the dual space is quite simple. In the following proposition, we show that every linear functional on a finite dimensional vector space can be completely characterised simply by knowing what values it takes at finitely many points (specifically, at the basis vectors), and that we can essentially cook up a linear functional to behave in ‘any way we want’ (i.e. take specific values at certain desired points).

Proposition 1.4.3. Let V be an n -dimensional vector space and let $\{v_1, \dots, v_n\}$ be a basis for V . Then:

- (a) Every linear functional $f \in V^*$ is completely determined by its values at v_1, \dots, v_n .
- (b) For any real numbers $a_1, \dots, a_n \in \mathbb{R}$ there exists an $f \in V^*$ such that $f(v_i) = a_i$ for all $i = 1, \dots, n$.

Proof. (a) Let $f \in V^*$ and let $x = k_1v_1 + \dots + k_nv_n \in V$. Then:

$$\begin{aligned} f(x) &= f(k_1v_1 + \dots + k_nv_n) \\ &= k_1f(v_1) + \dots + k_nf(v_n) \end{aligned} \quad (\text{by linearity of } f)$$

One sees from the above expression that $f(x)$ can be computed using only the knowledge of the values $f(v_1), \dots, f(v_n)$. Moreover, this *uniquely* identifies the linear functional. Suppose $g \in V^*$ agrees with f at each of the basis vectors, i.e. $f(v_i) = g(v_i)$ for all $i = 1, \dots, n$. Then for any $x \in V$ with representation as above, we have that:

$$\begin{aligned} g(x) &= g(k_1v_1 + \dots + k_nv_n) \\ &= k_1g(v_1) + \dots + k_ng(v_n) \\ &= k_1f(v_1) + \dots + k_nf(v_n) \\ &= f(x) \end{aligned}$$

(b) Let $a_1, \dots, a_n \in \mathbb{R}$. For $x = k_1v_1 + \dots + k_nv_n \in V$, define the mapping $f : V \rightarrow \mathbb{R}$ by:

$$f(x) = k_1a_1 + \dots + k_na_n$$

Then $f(v_i) = a_i$ for all $i = 1, \dots, n$, as required. We’ll now set about to show that f is linear.

Let $x \in V$ as above and let $y = l_1v_1 + \dots + l_nv_n \in V$. Then:

$$\begin{aligned} f(x + y) &= (k_1 + l_1)a_1 + \dots + (k_n + l_n)a_n \\ &= k_1a_1 + \dots + k_na_n + l_1a_1 + \dots + l_na_n \\ &= f(x) + f(y) \end{aligned}$$

Let $c \in \mathbb{R}$. Then:

$$\begin{aligned} f(cx) &= ck_1v_1 + \cdots + ck_nv_n \\ &= c(k_1v_1 + \cdots + k_nv_n) \\ &= cf(x) \end{aligned}$$

This proves that f is a linear map, so that $f \in V^*$. □

Given a vector space, how big is its dual space? In the infinite dimensional case, it can be shown via the axiom of choice that the dual space is always ‘bigger’ than the original space. In the finite dimensional case however, again we have it easy: the dimension of the dual space is the same as that of original space, so that they are isomorphic.

Theorem 1.4.4. *Let V be an n -dimensional vector space. Then its dual space V^* is also an n -dimensional vector space.*

Proof. We will prove that V^* is isomorphic to \mathbb{R}^n as vector spaces, which will imply that V^* has the same dimension as \mathbb{R}^n , which is n .

We first need to construct the function between V^* and \mathbb{R}^n which will serve as the candidate for what we hope will be the isomorphism. Fix a basis $\{v_1, \dots, v_n\}$ for V . Let $f \in V^*$. Define $\psi : V^* \rightarrow \mathbb{R}^n$ by:

$$\psi(f) = (f(v_1), \dots, f(v_n))^T$$

Let $f, g \in V^*$. Then:

$$\begin{aligned} \psi(f + g) &= (f(v_1) + g(v_1), \dots, f(v_n) + g(v_n))^T \\ &= (f(v_1), \dots, f(v_n))^T + (g(v_1), \dots, g(v_n))^T \\ &= \psi(f) + \psi(g) \end{aligned}$$

Let $a \in \mathbb{R}$. Then:

$$\begin{aligned} \psi(af) &= (af(v_1), \dots, af(v_n))^T \\ &= a(f(v_1), \dots, f(v_n))^T \\ &= a\psi(f) \end{aligned}$$

Thus, ψ is a linear map. We now proceed to show that ψ is a bijection.

Suppose $\psi(f) = \psi(g)$. Then $(f(v_1), \dots, f(v_n))^T = (g(v_1), \dots, g(v_n))^T$, i.e f and g agree at each of the basis vectors. By part (a) Proposition 1.4.3, this implies that $f = g$, thus proving the injectivity of ψ .

Now let $x = (a_1, \dots, a_n)^T \in \mathbb{R}^n$. By part (b) of Proposition 1.4.3, there exists an $f \in V^*$ such that $f(v_i) = a_i$ for all $i = 1, \dots, n$. It follows that $\psi(f) = x$. Since the choice of $x \in \mathbb{R}^n$ was arbitrary, it follows that $\mathcal{R}(\psi) = \mathbb{R}^n$, thus proving surjectivity of ψ . Putting all of this together shows that ψ is a bijective linear function, i.e a vector space isomorphism. □

Let V be an n -dimensional vector space. Since the dual space V^* is n -dimensional, it has a basis consisting of n elements. Given a basis $\{v_1, \dots, v_n\}$ for V , there is a particular nice choice for a basis on V^* . For each $i = 1, \dots, n$, define the linear functional $v_i^* : V \rightarrow \mathbb{R}$ by:

$$v_i^*(v_j) = \delta_{ij}, \quad \text{for } j = 1, \dots, n$$

So the linear functionals v_i^* satisfy a ‘bi-orthogonality property’. The fact that such linear functionals exist again comes from Proposition 1.4.3. The set $\{v_1^*, \dots, v_n^*\}$ is a linearly independent set in V^* . To see

this, let $x = a_1v_1 + \cdots + a_nv_n \in V$ be arbitrary and consider the following vector equation:

$$\begin{aligned}
 0 &= \sum_{i=1}^n c_i v_i^*(x) \\
 &= \sum_{i=1}^n c_i v_i^* \left(\sum_{j=1}^n a_j v_j \right) \\
 &= \sum_{i=1}^n \sum_{j=1}^n c_i a_j v_i^*(v_j) \\
 &= \sum_{i=1}^n \sum_{j=1}^n c_i a_j \delta_{ij} \\
 &= \sum_{i=1}^n c_i a_i
 \end{aligned}$$

Since the above equation must hold for all $x \in V$, choosing $x = v_k$ for each $k = 1, \dots, n$ means that $a_i = \delta_{ik}$ in the above expression, which gives $c_k = 0$. This demonstrates the linear independence of the v_i^* 's. Since we have shown that V^* is an n -dimensional vector space, and we have that $\{v_1^*, \dots, v_n^*\}$ is a linearly independent set in V^* consisting of n elements, we know automatically that this must be a basis for V^* . We call this special basis the **dual basis**.

Now, since the dual space V^* is itself a vector space, we can consider its dual space V^{**} , called the **second dual**. We know that V^{**} must also be a vector space with the same dimension as V , so they are isomorphic. Better yet, there is a natural identification of elements from V with elements from V^{**} , so that the isomorphism is canonical.

Theorem 1.4.5. *Let V be an n -dimensional vector space. Then V is canonically isomorphic to its second dual V^{**} .*

Proof. For each $x \in V$, define the mapping $\phi_x : V^* \rightarrow \mathbb{R}$ by:

$$\phi_x(f) = f(x)$$

i.e ϕ_x is an evaluation map. Given a linear functional on V^* , it returns a scalar by evaluating it at the fixed point x . We will show that ϕ_x is a linear map.

Let $f, g \in V^*$. Then $\phi_x(f + g) = (f + g)(x) = f(x) + g(x) = \phi_x(f) + \phi_x(g)$.

Let $a \in \mathbb{R}$. Then $\phi_x(af) = af(x) = a\phi_x(f)$.

Thus, ϕ_x is a linear functional on V^* , so $\phi_x \in V^{**}$. We can hence define the mapping $\psi : V \rightarrow V^{**}$ by $\psi(x) = \phi_x$. We will show that ψ is a vector space homomorphism. Let $x, y \in V$ and let $f \in V^*$. Then:

$$\begin{aligned}
 (\psi(x + y))(f) &= \phi_{x+y}(f) \\
 &= f(x + y) \\
 &= f(x) + f(y) \\
 &= \phi_x(f) + \phi_y(f) \\
 &= (\psi(x))(f) + (\psi(y))(f)
 \end{aligned}$$

Since the choice of $f \in V^*$ was arbitrary, this shows that $\psi(x + y) = \psi(x) + \psi(y)$.

Let $a \in \mathbb{R}$. Then:

$$\begin{aligned}
 (\psi(ax))(f) &= \phi_{ax}(f) \\
 &= f(ax) \\
 &= af(x) \\
 &= a\phi_x(f) \\
 &= a(\psi(x))(f)
 \end{aligned}$$

Thus, $\psi(ax) = a\psi(x)$. This proves that ψ is a vector space homomorphism.

To show the injectivity of ψ , we will prove that its kernel is trivial. Suppose $x \in \ker \psi$, so $\psi(x) = 0$. This means that for all $f \in V^*$, $f(x) = 0$. If $x \neq 0$, then with respect to some basis $\{v_1, \dots, v_n\}$ of V , x can be written as a nontrivial linear combination of the basis vectors: $x = a_1v_1 + \dots + a_nv_n$, where at least one of the a_i 's are nonzero. Pick a nonzero coefficient and call it a_k . Then by part (b) of Proposition 1.4.3, there exists a $g \in V^*$ such that $g(v_k) = 1$ and $g(v_i) = 0$ for all $i \neq k$. Then:

$$\begin{aligned}
 g(x) &= g(a_1v_1 + \dots + a_nv_n) \\
 &= a_1g(v_1) + \dots + a_kg(v_k) + \dots + a_ng(v_n) \\
 &= a_k \\
 &\neq 0
 \end{aligned}$$

This is a contradiction. Hence, $x = 0$, which shows that $\ker \psi = \{0\}$, and hence ψ is injective.

Finally, we show that ψ is surjective. Let $\chi \in V^{**}$. Let $\{v_1^*, \dots, v_n^*\}$ be the dual basis for $\{v_1, \dots, v_n\}$. Take $a_i = \chi(v_i^*)$ for $i = 1, \dots, n$, and thus construct the element $x = a_1v_1 + \dots + a_nv_n \in V$. We claim that $\chi = \phi_x$. Let $f \in V^*$, and expand it in terms of the dual basis:

$$f = \sum_{i=1}^n c_i v_i^*$$

Then we have:

$$\begin{aligned}
 \phi_x(f) &= f(x) \\
 &= \sum_{i=1}^n c_i v_i^*(x) \\
 &= \sum_{i=1}^n c_i v_i^* \left(\sum_{j=1}^n a_j v_j \right) \\
 &= \sum_{i=1}^n \sum_{j=1}^n c_i a_j v_i^*(v_j) \\
 &= \sum_{i=1}^n \sum_{j=1}^n c_i a_j \delta_{ij} \\
 &= \sum_{i=1}^n c_i a_i \\
 &= \sum_{i=1}^n c_i \chi(v_i^*) \\
 &= \chi \left(\sum_{i=1}^n c_i v_i^* \right) && \text{(by linearity of } \chi \text{)} \\
 &= \chi(f)
 \end{aligned}$$

The above argument holds for all $f \in V^*$, and thus we conclude that $\chi = \phi_x = \psi(x)$. This proves that ψ is surjective, and hence a bijection, and hence an isomorphism. \square

The property that a vector space is naturally isomorphic to its second dual is referred to as *algebraic reflexivity*. This is another property that is guaranteed to hold in the finite dimensional setting, but not in the infinite dimensional case. There is always a natural monomorphism (injective homomorphism) from a vector space to its second dual, however it is an isomorphism if and only if the vector space is finite dimensional.

Naturally, we can keep playing the game of successively finding the dual space of a dual space, but the results in this section show that this is not necessary. It turns out in the finite dimensional setting that a vector space and its dual are of equal dimension, and are isomorphic. So the study of the dual space of a finite dimensional vector space reduces to the study of the vector space itself.

1.5 Topology of Euclidean Space

So far, we have mostly concerned ourselves with the algebraic properties of Euclidean space. We will now consider the structure of Euclidean space that makes it amiable to do calculus on. Recall that the Euclidean inner product induces a vector norm, which in turn induces a metric. To add to this already lengthy chain of implications, the metric induces a *topology* on our Euclidean space. A topological space is the most general setting in which one can define the notion of continuity, which is certainly something we would like to have in a calculus setting!

Since the topology of Euclidean space is the metric topology (where the metric is the Euclidean metric d_2), we will define the open sets via the notion of *open balls*. We have already described the Euclidean metric for \mathbb{R}^n . To generalise it to Euclidean n -space E , take an orthonormal basis $\{e_1, \dots, e_n\}$ of E . Then for any $x = x_1e_1 + \dots + x_ne_n \in E$, the Euclidean norm is given by:

$$\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$$

with corresponding Euclidean metric:

$$d_2(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

where $y = y_1e_1 + \dots + y_ne_n$.

Definition 1.5.1. Let E be a Euclidean n -space. An **open ball** of centre $x \in E$ and radius $r > 0$ is the set $B(x, r) = \{y \in E \mid d_2(x, y) < r\} = \{y \in E \mid \|x - y\| < r\}$.

The **closed ball** of centre x and radius r is the set $\overline{B}(x, r) = \{y \in E \mid d_2(x, y) \leq r\} = \{y \in E \mid \|x - y\| \leq r\}$.

An open ball centered at a point $x \in E$ with radius r simply consists of all of the points that are within a distance r of x . The closed ball also includes those points that are exactly a distance of r away from x (as measured using the Euclidean metric).

Definition 1.5.2. Let E be a Euclidean n -space. A set $U \subset E$ is said to be **open** if for every $x \in U$, there exists an $r > 0$ such that $B(x, r) \subset U$.

Some trivial examples of open sets include the empty set, and the whole space E . Open balls themselves are also open sets. An open rectangle, which is simply the Cartesian product of open intervals in \mathbb{R} , is an open set in \mathbb{R}^n .

Now we address the age-old question of: ‘once we have a bunch of open sets, how do we make new open sets from old?’ This is answered by:

Proposition 1.5.3. *Let E be a Euclidean n -space. Then:*

(a) *If $U_\alpha \subset E$ is open for each $\alpha \in A$, then the union*

$$U = \bigcup_{\alpha \in A} U_\alpha$$

is open in E .

(b) *If U_1, \dots, U_k are open subsets of E , then the intersection*

$$U = U_1 \cap \dots \cap U_k$$

is open in E .

Proof. (a) Let $U_\alpha \subset E$ be open for each $\alpha \in A$, and consider the set $U = \bigcup_{\alpha \in A} U_\alpha$. Let $x \in U$. Then there exists a $\beta \in A$ such that $x \in U_\beta$. Since U_β is open, there exists an $r > 0$ such that $B(x, r) \subset U_\beta \subset U$. Thus, U is open.

(b) Let U_1, \dots, U_k be open, and consider the set $U = U_1 \cap \dots \cap U_k$. Let $x \in U$. For each $i = 1, \dots, k$, $x \in U_i$, so there exists an $r_i > 0$ such that $B(x, r_i) \subset U_i$. Choose $r = \min \{r_1, \dots, r_k\} > 0$. Then $B(x, r) \subset U_i$ for all $i = 1, \dots, k$. Thus, $B(x, r) \subset U$, and hence U is open. \square

So arbitrary unions of open sets are open, and *finite* intersections of open sets are open. Infinite intersections of open sets are not necessarily open. For example, consider $I_n = (-\frac{1}{n}, \frac{1}{n})$, which is an open set in \mathbb{R} for each $n \in \mathbb{N}$. But:

$$\bigcap_{n=1}^{\infty} I_n = \{0\}$$

which is *not* open in \mathbb{R} .

Proposition 1.5.3 above, in combination with the fact that \emptyset and E are open, shows that the collection of open sets in a Euclidean space E forms a *topology* on E . This fact did not appeal to anything particular about Euclidean space other than its underlying metric space structure.

We will later want to define the notions of limits and continuity, which both involve the idea of points being forced to being ‘really close to each other’. Frequently in analysis we will also allow for our variables to vary in a ‘sufficiently small amount’. To make formal these ideas, we use the topological concept of ‘closeness’ in terms of *neighbourhoods* of points:

Definition 1.5.4. Let E be a Euclidean n -space and let $x \in E$. A **neighbourhood** of x is a subset $U \subset E$ such that there exists an open subset $V \subset U$ that contains x .

Note that in the metric topology, an open set containing a point $x \in E$ must contain an open ball $B(x, r)$ for some $r > 0$, so that we can also state our definition of a neighbourhood of a point x as simply ‘a set which contains an open ball centred at x ’. Hence, the topological notion of ‘wiggling around x in an arbitrarily small amount’ (constrained the variation to smaller and smaller neighbourhoods) equates to the metric notion of small movement (constrained the variation to smaller and smaller open balls).

Thus, an alternative way to characterise open sets is the statement that ‘a set is open if and only if it is a neighbourhood of all of its points’, which follows immediately from the definition. We can also rephrase this in terms of *interior points*, which we define below:

Definition 1.5.5. Let E be a Euclidean n -space. An **interior point** of a set $S \subset E$ is an $x \in S$ such that there exists an $r > 0$ such that $B(x, r) \subset S$.

One sees immediately from the definition above and from the definition of an open set that a set is open if and only if all of its points are interior points. So if a set S has any points that are not interior points, it cannot be open. However, if we consider the subset of S comprising of only the interior points of S , then we get an open subset of S . In fact, this is the largest open subset contained in S , which we call the *interior* of S .

Definition 1.5.6. Let E be a Euclidean n -space and let $S \subset E$. The **interior** of S , denoted S° , is the union of all open subsets of S .

Our definition of the interior above formulates it as the largest open subset contained in S , but it can be readily shown that the interior is simply the set of interior points of S . This means that a set is open if and only if it is equal to its interior.

Also equally important in a topological space is the notion of a *closed* set:

Definition 1.5.7. Let E be a Euclidean n -space. A set $C \subset E$ is said to be **closed** if it is the complement of an open set, i.e. $C = E \setminus U$ for some open set $U \subset E$.

One should be careful that a set being closed is *not* the negation of a set being open. There are sets that are neither open nor closed (e.g. the half-open interval in \mathbb{R} : $[a, b)$), and sets that are *both* open and closed (the empty set and the entire space).

Closed balls are a simple example of closed sets. Cartesian products of closed intervals in \mathbb{R} are closed sets in \mathbb{R}^n . An important example of closed sets in Euclidean space are vector subspaces. Complementary to Proposition 1.5.3, we now list the ways one can combine closed sets to make new closed sets:

Proposition 1.5.8. *Let E be a Euclidean n -space. Then:*

(a) *If $C_\alpha \subset E$ is closed for each $\alpha \in A$, then the intersection*

$$C = \bigcap_{\alpha \in A} C_\alpha$$

is closed in E .

(b) *If C_1, \dots, C_k are closed subsets of E , then the union*

$$C = C_1 \cup \dots \cup C_k$$

is closed in E .

Proof. (a) Let $C_\alpha \subset E$ be closed for each $\alpha \in A$. Then $E \setminus C_\alpha$ is open for all $\alpha \in A$. Hence, $\bigcup_{\alpha \in A} (E \setminus C_\alpha)$ is an open set, since arbitrary unions of open sets are open. But by de Morgan's Laws, we have that:

$$\bigcup_{\alpha \in A} (E \setminus C_\alpha) = E \setminus \left(\bigcap_{\alpha \in A} C_\alpha \right)$$

We thus see that $C = \bigcap_{\alpha \in A} C_\alpha$ is closed.

(b) Let C_1, \dots, C_k be closed. Then $E \setminus C_i$ is open for all $i = 1, \dots, k$. Thus, $(E \setminus C_1) \cap \dots \cap (E \setminus C_k)$ is open since finite intersections of open sets are open. But by de Morgan's Laws, we have that:

$$(E \setminus C_1) \cap \dots \cap (E \setminus C_k) = E \setminus (C_1 \cup \dots \cup C_k)$$

It immediately follows that $C_1 \cup \dots \cup C_k$ is closed. □

One way to study closed sets, and the degree at which a set fails to be closed is via the notion of *limit points*:

Definition 1.5.9. Let E be a Euclidean n -space and let $S \subset E$. A **limit point** of S is a point $x \in E$ such that for every $r > 0$, $S \cap (B(x, r) \setminus \{x\}) \neq \emptyset$.

The motivation for the name 'limit point' arises from the fact you can approach a limit point of a set and get arbitrarily close to it even whilst remaining completely within the set. That is, it is the limit of a nonconstant sequence consisting entirely of points from the set.

One important fact about closed sets is that the limit of a sequence consisting of points from the set must stay inside the set. From the above discussion, it is then clear that a closed set must contain all of its limit points. This turns out to be both a necessary and sufficient criteria for a set to be closed.

Proposition 1.5.10. *Let E be a Euclidean n -space. A set is closed in E if and only if it contains all of its limit points.*

Proof. (\implies) Suppose $S \subset E$ is closed, but assume for a contradiction that S does not contain all of its limit points. Let $x \in E$ be a limit point of S but $x \notin S$. Then $x \in E \setminus S$. However, we know that $E \setminus S$ is an open set (since S is closed), so there exists an $r > 0$ such that $B(x, r) \subset E \setminus S$. But since x is a limit point of S , then there exists a $y \in S \cap (B(x, r) \setminus \{x\})$, so that $y \in S \cap E \setminus S$, which is impossible. Thus, S must contain all of its limit points.

(\impliedby) Suppose S contains all of its limit points. Let $x \in E \setminus S$, so that x is not a limit point of S . Therefore, there exists an $r > 0$ such that $S \cap (B(x, r) \setminus \{x\}) = \emptyset$. Hence, $B(x, r) \subset E \setminus S$. This proves that $E \setminus S$ is open, and hence that S is closed. \square

Given a set S that is not necessarily closed, we can ‘add’ points to it until we obtain a closed set. The smallest set that results after we add the minimum amount of points required to make the set closed is called the closure of the set:

Definition 1.5.11. Let E be a Euclidean n -space and let $S \subset E$. The **closure** of S , denoted \overline{S} , is the intersection of all closed sets containing S .

It is clear from the above definition that the closure of $S \subset E$ is the smallest closed set containing S , consistent with our above discussion. In fact, the additional points that we need to add to S to make it closed are simply the limit points of S which it does not already contain. Hence, the closure of S is simply the union of S with its set of limit points.

1.6 Completeness

1.7 Continuity

1.8 Compactness

1.9 Equivalence of Norms

1.10 The Space $L(E, F)$